

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
24. Oktober 2002 (24.10.2002)

PCT

(10) Internationale Veröffentlichungsnummer  
**WO 02/084644 A1**

(51) Internationale Patentklassifikation<sup>7</sup>: G10L 19/00, 11/02 (72) Erfinder; und  
(75) Erfinder/Anmelder (nur für US): BERGER, Jens [DE/DE]; Jablonskistrasse 23, 10405 Berlin (DE).

(21) Internationales Aktenzeichen: PCT/DE02/01200 (74) Gemeinsamer Vertreter: DEUTSCHE TELEKOM AG; Rechtsabteilung (Patente) PA1, 64307 Darmstadt (DE).

(22) Internationales Anmeldedatum: 3. April 2002 (03.04.2002) (81) Bestimmungsstaat (national): US.

(25) Einreichungssprache: Deutsch (84) Bestimmungsstaaten (regional): europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

(26) Veröffentlichungssprache: Deutsch

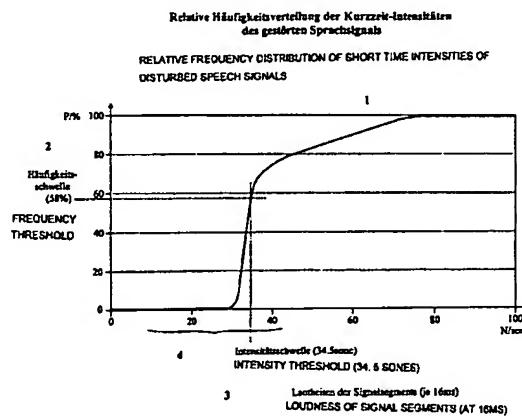
(30) Angaben zur Priorität: 101 20 168.0 18. April 2001 (18.04.2001) DE Veröffentlicht:  
— mit internationalem Recherchenbericht  
— vor Ablauf der für Änderungen der Ansprüche geltenden Frist; Veröffentlichung wird wiederholt, falls Änderungen eintreffen

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): DEUTSCHE TELEKOM AG [DE/DE]; Friedrich-Ebert-Allee 140, 53113 Bonn (DE).

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR DETERMINING INTENSITY PARAMETERS OF BACKGROUND NOISE IN SPEECH PAUSES OF VOICE SIGNALS

(54) Bezeichnung: VERFAHREN ZUR BESTIMMUNG VON INTENSITÄTSKENNWERTEN VON HINTERGRUNGERÄUSCHEN IN SPRACHPAUSEN VON SPRACHSIGNALEN



WO 02/084644 A1

(57) Abstract: Known methods for determining intensity parameters are based on the evaluation of short signal segments and their direct allocation to speech pauses or speech activity. In order to distinguish speech from speech pauses, intensity thresholds are often used. When the undisturbed source signal is used to mark speech pauses, a variably occurring time lag between source voice signal and disturbed voice signal often impedes exact transfer of the marking. Intensity parameters of background noises in speech pauses can be determined from the frequency distribution of the intensity values for short signal segments using the method disclosed in the invention. In order to assign intensity values, the fraction of speech pauses in the entire signal is calculated from the undisturbed source signal and defined as frequency threshold. Intensity values below the frequency threshold are assigned to the speech pauses. The arithmetic mean value of said intensity value is determined as intensity parameter for the background noise in the speech pauses. Percentile parameters for background noises in speech pauses can also be calculated with the inventive method.

[Fortsetzung auf der nächsten Seite]

BEST AVAILABLE COPY



Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

---

(57) **Zusammenfassung:** Bekannte Verfahren zur Bestimmung von Intensitätskennwerten basieren auf der Auswertung von kurzen Signalsegmenten und deren direkter Zuordnung zu Sprachpausen bzw. zur Sprachaktivität. Zur Unterscheidung von Sprache und Sprachpausen werden häufig Intensitätsschwellen benutzt. Wird das ungestörte Quellsignal zur Markierung der Sprachpausen verwendet, lässt häufig ein auftretender variabler Zeitversatz zwischen Quellsprachsignal und gestörtem Sprachsignal eine exakte Übernahme der Markierung nicht zu. Mit dem vorliegenden Verfahren werden Intensitätskennwerte von Hintergrundgeräuschen in den Sprachpausen aus der Häufigkeitsverteilung der Intensitätswerte für kurze Signalsegmente ermittelt. Für die Zuordnung der Intensitätswerte wird der Anteil an Sprachpausen im Gesamtsignal aus dem ungestörten Quellsignal berechnet und als Häufigkeitsschwelle definiert. Die Intensitätswerte unterhalb der Häufigkeitsschwelle werden den Sprachpausen zugeordnet. Der arithmetische Mittelwert dieser Intensitätswerte wird als Intensitätskennwert für die Hintergrundgeräusche in den Sprachpausen bestimmt. Mit dem vorliegenden Verfahren können auch Perzentil-Kennwerte für die Hintergrundgeräusche in den Sprachpausen berechnet werden.

## **Verfahren zur Bestimmung von Intensitätskennwerten von Hintergrundgeräuschen in Sprachpausen von Sprachsignalen**

### **Beschreibung**

5

#### ***Vorbemerkung***

Die Erfindung bezieht sich auf ein Verfahren zur Bewertung von Hintergrundgeräuschen in Sprachpausen von aufgezeichneten oder übertragenen Sprachsignalen.

10 Die empfundene Sprachqualität, z.B. in Telefonverbindungen oder Rundfunkübertragungen, wird hauptsächlich von sprachsimultanen Störungen, also von Störungen während der Sprachaktivität, bestimmt. Aber auch Geräusche in den Sprachpausen gehen in das Qualitätsurteil ein, insbesondere bei hochqualitativer Sprachwiedergabe.

15 Die Intensität des Hintergrundgeräusches in den Sprachpausen kann als ergänzender Kennwert zur Bestimmung der Sprachqualität (Sprachgüte) verwendet werden.

Sprachqualitätsbestimmungen von Sprachsignalen werden in der Regel mittels auditiver ("subjektiver") Untersuchungen mit Versuchspersonen vorgenommen.

20 Das Ziel von instrumentellen ("objektiven") Verfahren zur Sprachqualitätsbestimmung ist es dagegen, aus Eigenschaften des zu bewertenden Sprachsignals mittels geeigneter Rechenverfahren Kennwerte zu ermitteln, die die Sprachqualität des Sprachsignals beschreiben, ohne auf Urteile von Versuchspersonen zurückgreifen zu müssen.

25 Eine sichere Qualitätsbestimmung liefern instrumentelle Verfahren, die auf einem Vergleich von ungestörtem Referenzsprachsignal (Quellsprachsignal) und dem gestörten Sprachsignal am Ende der Übertragungskette beruhen. Es existieren viele solcher Verfahren, die meist in sogenannten Probeverbundungssystemen eingesetzt werden. Dabei wird an der Quelle das ungestörte Quellsprachsignal eingespeist und nach der Übertragung wieder aufgezeichnet.

30

***Stand der Technik und Nachteile bekannter Verfahren***

Bekannte Verfahren zur Bestimmung der Intensität von Hintergrundgeräuschen gehen

meist vom gestörten Signal selbst aus und nutzen eine festgelegte Intensitätsschwelle zur

5 Unterscheidung von aktiver Sprache und Sprachpausen (Fig. 1). Diese Schwelle ist im einfachsten Fall konstant im Verfahren eingestellt, kann aber auch anhand des Signalverlaufs adaptiert werden (z.B. festgelegter Abstand zum Signal-Spitzenwert). Das Ziel ist eine sichere Unterscheidung zwischen Sprache und Sprachpause. Gelingt die Unterscheidung, können die gesuchten Intensitätskennwerte des Hintergrundgeräusches aus den 10 als Sprachpause detektierten Signalabschnitten bestimmt werden. Dazu werden im Allgemeinen die als Sprachpause detektierten Signalabschnitte nochmals in kürzere Segmente (typisch sind 8...40ms) unterteilt und für diese die Intensitätsberechnungen (z.B. Effektivwert oder Lautheit) vorgenommen. Aus den Ergebnissen können dann 15 Intensitätskennwerte bestimmt werden.

15 Die Verfahren liefern bei geringen Geräuschintensitäten in Sprachpausen und gleichzeitig hoher Intensität der Sprache (großes Sprach-Geräusch-Verhältnis) sichere Meßwerte, da die Unterscheidung zwischen Sprache und Sprachpause sicher erfolgen kann (Fig.1).

20 Bei steigenden Geräuschintensitäten in Sprachpausen (abnehmendes Sprach-Geräusch-Verhältnis) treten zunehmend Unsicherheiten in der Unterscheidung zwischen Sprache und Sprachpausen auf. Hier ist es schwierig den Schwellenwert so festzulegen, dass zum einen keine Geräuschabschnitte mit höheren Intensitäten als Sprache detektiert werden (Schwelle zu niedrig) und zum anderen keine Sprachabschnitte geringerer Intensität als Sprachpause 25 gewertet werden (Schwelle zu hoch) (Fig. 2).

Erreicht die Intensität des Geräusches in den Sprachpausen die Intensität der aktiven Sprache oder übersteigt diese sogar, ist keine Intensitätsschwelle zu finden, die eine Unterscheidung zwischen Sprache und Sprachpause ermöglicht.

Lösungen für die beschriebenen Probleme sind möglich, wenn z.B. unterschiedliche spektrale Charakteristika von Sprache und Hintergrundgeräuschen vorliegen. Hier kann durch geeignete Vorfilterung des Signals bzw. durch eine spektrale Analyse und Auswertung von ausgewählten Frequenzbändern ein höheres Verhältnis von Sprache zu

5 Hintergrundgeräusch in den betrachteten Frequenzbereichen erreicht werden, so dass wieder eine sichere Unterscheidung zwischen aktiver Sprache und Sprachpause möglich ist.

Andere Lösungen bedienen sich bestimmter Parameter, die bei Sprachcodierung ermittelt

10 werden und nutzen diese zur Unterscheidung zwischen Sprache und Abschnitten mit Hintergrundgeräuschen. Dabei ist es das Ziel, aus den Parametern abzuleiten, ob das betrachtete Signalsegment typische Eigenschaften von Sprache (z.B. stimmhafte Anteile) aufweist. Ein Beispiel hierfür ist "Voice-Activity Detector" (ETSI Recommendation GSM 06.92, Valboune, 1989).

15 Diese Verfahren arbeiten bei geringen Sprach-Geräusch-Verhältnissen robuster und werden vorrangig zur Unterdrückung der Übertragung von Sprachpausen z.B. im Mobilfunk eingesetzt. Die Verfahren zeigen jedoch Unsicherheiten, wenn das Hintergrundgeräusch selbst Sprache beinhaltet oder sprachähnlich ist. Solche Abschnitte werden dann als

20 Sprache klassifiziert, obwohl sie von einem Zuhörer als störendes Hintergrundgeräusch empfunden werden.

Instrumentelle Sprachqualitätsmessverfahren basieren meist auf dem Prinzip des Signalvergleichs von ungestörtem Referenzsprachsignal und gestörtem und zu bewertenden

25 Signal. Beispiele hierfür sind die Veröffentlichungen:

"A perceptual speech-quality measure based on a psychacoustic sound representation"  
(Beerends, J. G.; Stemerdink, J. A., J. Audio Eng. Soc. 42(1994)3, S.115-123)

30 "Auditory distortion measure for speech coding" (Wang, S; Sekey, A.; Gersho, A.: IEEE Proc. Int. Conf. acoust., speech and signalprocessing (1991), S.493-496).

Der derzeit gültige ITU-T Standard P.861 beschreibt ebenfalls ein derartiges Verfahren: "Objective quality measurement of telephone-band speech codecs" (ITU-T Rec. P.861, Genf 1996).

5

Solche Messverfahren werden in sogenannten Probeverbindungssystemen eingesetzt, bei denen ein bekanntes Referenzsprachsignal (Quellsprachsignal) an der Quelle eingespeist, über z. B. eine Telefonverbindung übertragen und an der Senke aufgezeichnet wird. Nach der Aufzeichnung des Sprachsignals werden zur Bewertung der Sprachqualität des möglicherweise gestörten Signals dessen Eigenschaften mit denen des ungestörten Quellsprachsignals verglichen.

Steht für die Bestimmung des Hintergrundgeräusches in Sprachpausen das ungestörte Quellsprachsignal zur Verfügung, dann kann dieses zur Festlegung der Übergangszeitpunkte von Sprache zur Sprachpause bzw. von Sprachpause zur Sprache benutzt werden. Dazu wird z.B. ein Verfahren mit Schwellwertbestimmung - wie oben beschrieben - auf das Quellsprachsignal angewandt. Das Verfahren liefert sichere Unterscheidungen zwischen Sprache und Sprachpause, da das Sprach-Geräusch-Verhältnis im ungestörten Quellsprachsignal ausreichend hoch ist (Fig. 3a). Die Zeitpunkte der Schwellpassage, d.h. Beginn bzw. Ende der Sprachaktivität, können nun auf das gestörte Sprachsignal übertragen werden (Fig. 3b).

Unproblematisch kann ein solches Verfahren modifiziert werden, wenn zwischen Quellsprachsignal und gestörtem Signal eine konstante Zeitdifferenz (z.B. Verzögerung durch Signalübertragung) eintritt. Bedingung ist aber, dass diese Zeitdifferenz vorab sicher bestimmt werden kann und dann zur Korrektur der Zeitpunkte Ende bzw. Beginn der Sprachaktivität genutzt wird. Das ist meist bei zeit-invarianten Systemen möglich, da diese eine konstante Verzögerung besitzen (Fig. 3c). Prinzipiell funktioniert ein solches Verfahren auch, wenn der Zeitversatz zwischen beiden Signalen nicht für die gesamte Signallänge konstant ist, sondern variabel verläuft. Zu diesen zeit-invarianten Systemen zählen insbesondere paket-basierte Übertragungssysteme,

bei denen durch unterschiedliche Paketlaufzeiten und entsprechendes Management im Empfänger deutliche Schwankungen in der Systemverzögerung auftreten können. Um Verlusten durch verspätet eintreffende Pakete vorzubeugen, werden teilweise Sprachpausen im Empfänger verlängert und spätere wieder verkürzt. Eine Übertragung der

5 Zeitpunkte von Beginn bzw. Ende der Sprachaktivität ist nur noch bei Kenntnis der aktuellen Verzögerung an diesen Punkten möglich. Die adaptive Bestimmung des Zeitversatzes ist rechenzeitintensiv und gelingt insbesondere bei verringerten Sprach-Geräusch-Verhältnissen oft nur unzureichend. Wenn die adaptive Bestimmung des Zeitversatzes nicht sicher gelingt, können Anfang und Ende von Sprachpausen nicht exakt  
10 oder gar nicht ermittelt werden. Dadurch ist keine oder nur eine unsichere Bestimmung der Intensitätskennwerte von Pausengeräuschen möglich.

### **Aufgabe**

Wie beschrieben, ist die Bestimmung von Hintergrundgeräuschen in Sprachpausen auch  
15 bei Kenntnis des ungestörten Quellsprachsignals schwierig oder teilweise unmöglich, insbesondere wenn

- ein geringes Verhältnis von Sprache zu Hintergrundgeräusch vorliegt,
- das Hintergrundgeräusch Sprache beinhaltet oder selbst sprachähnlich ist,
- der Zeitversatz zwischen ungestörtem Quellsprachsignal und gestörtem Sprachsignal  
20 nicht konstant über die gesamte Signallänge ist.

Es soll ein Verfahren vorgestellt werden, mit dem auch unter den genannten Bedingungen eine sichere und schnelle Bestimmung von Intensitätskennwerten des Hintergrundgeräusches in Sprachpausen gewährleistet wird. Bedingung ist, dass sowohl Quellsprachsignal als auch gestörtes Sprachsignal vollständig aufgezeichnet zur Verfügung stehen.  
25

### **Lösungsprinzip**

Die bekannten Verfahren gehen davon aus, den Zeitpunkt von Beginn und Ende einer Sprachpause möglichst exakt zu ermitteln. Im Ergebnis steht dann das Signal von den  
30 Pausenabschnitten zur weiteren Auswertung zur Verfügung. Aus diesen separierten Pausenabschnitten des Signals werden die Intensitätskennwerte ermittelt.

Mit dem vorliegenden Verfahren können Intensitätskennwerte von Hintergrundgeräuschen in Sprachpausen von Sprachsignalen bestimmt werden, ohne dass die exakten Zeitpunkte von Beginn und Ende eines Pausenabschnitts ermitteln werden müssen. Auch ist eine

5 Separierung des Sprachpausensignals für die Auswertung nicht erforderlich.

Basis für das hier beschriebene Verfahren zur Bestimmung von Intensitätskennwerten von Hintergrundgeräuschen in Sprachpausen von Sprachsignalen ist die kumulative Häufigkeitsverteilung der Intensitätswerte von den Signalsegmenten, in die das Sprachsignal

10 zuvor unterteilt wird. Diese Kurzzeit-Signalintensitäten beziehen sich auf Signalsegmente mit einer Dauer von z.B. 8ms oder 16ms. Die Häufigkeitsverteilung gibt an, wie hoch der Anteil an Kurzzeit-Intensitäten unterhalb eines definierten Schwellwertes ist.

Für die Berechnung der Häufigkeitsverteilung wird das zu analysierende Sprachsignal in

15 kurze aufeinanderfolgende Signalsegmente unterteilt und von jedem Signalsegment der Intensitätswert (z.B. Lautheit oder Effektivwert) bestimmt.

Fig. 4 zeigt einen typischen Kurvenverlauf für Sprachsignale mit stationärem Hintergrundgeräusch (Sprach-Geräusch-Abstand ca. 10dB). Die kumulative Häufigkeitsverteilung ist

20 am Beispiel von Kurzzeit-Lautheiten (Lautheiten berechnet nach ISO532) dargestellt. Ausgewertet wurden 2000 Segmente von 16ms Länge. Es ist zu erkennen, dass keines der Segmente einen geringeren Wert als 30 sone aufweist ( $P = 0\%$ ) und auch kein Segment eine höhere Lautheit als 80 sone erreicht, da hier schon der Wert  $P=100\%$  erreicht wird. Der steile Anstieg der Funktion bei ca. 30 sone lässt auf eine geringe Fluktuation der

25 Signalintensität in großen Bereichen (fast 70%) des Signals schließen. Als Signal wurde hier ein Sprachsignal mit additiven weißen Rauschen benutzt.

Eine solche Verteilungsfunktion soll nun dazu benutzt werden, Intensitätskennwerte von Hintergrundgeräuschen in den Sprachpausen zu ermitteln. Dazu ist es erforderlich, den

30 Anteil an Sprachpausen im Gesamtsignal zu kennen. Dieser Anteil kann aus dem ungestörten Quellsprachsignal bestimmt werden (Fig. 3a).

$$\text{Gesamtlänge der Sprachpausen} = (t1 - t0) + (t3 - t2)$$

$$\text{Gesamtlänge des Signalabschnitts} = (t4 - t0)$$

$$\text{Sprachpausenanteil} = \frac{\text{Gesamtlänge der Sprachpausen}}{\text{Gesamtlänge des Signalabschnitts}}$$

5 Wird davon ausgegangen, dass das Verhältnis von aktiver Sprache zu Sprachpausen während der Übertragung weitgehend konstant bleibt, kann dieser Wert auch auf das gestörte Signal übertragen werden.

Ist der Anteil an Sprachpausen am gesamten Sprachsignal bekannt und wird dieser Anteil  
10 als Häufigkeitsschwelle definiert, so kann aus der Häufigkeitsverteilung der Kurzzeit-Intensitäten der Häufigkeitsschwelle entsprechende Intensitätsschwellwert ermittelt werden.

In Fig. 4 ist als Beispiel ein Anteil an Sprachpausen von 58 % eingetragen. Dieser  
15 Häufigkeitsschwelle  $P_z = 0.58$  entspricht ein Intensitätsschwellwert von  $N = 34.5 \text{ sone}$ , das bedeutet, dass von 58 % der Signalsegmente der Intensitätswert (Lautheit) von 34,5 sone nicht überschritten wird.

Der Bereich unterhalb des Intensitätsschwellwertes zeigt die Häufigkeitsverteilung für  
20 Intensitätswerte von Signalsegmenten in den Sprachpausen und kann für die Ermittlung von Intensitätskennwerten von den Hintergrundgeräuschen in den Sprachpausen benutzt werden.

Es wird davon ausgegangen, dass kein Sprachpausensegment einen höheren Intensitätswert  
25 als ein Sprachsegment besitzt, so dass der Intensitätsschwellwert als Maximalwert für das Hintergrundgeräusch in Sprachpausen angesehen werden kann.

**Ermittlung des arithmetischen Mittelwertes von Intensitäten**

Aus der kumulativen Verteilungsfunktion lässt sich auch der arithmetische Mittelwert aller Segmente ableiten, deren Intensitäten sich unter einer vorher ermittelten Häufigkeitsschwelle befinden. Dazu ist zunächst eine Differenzierung der kumulativen Verteilungsfunktion  $P(x)$  in eine Verteilungsdichtefunktion  $p(x)$  vorzunehmen.

Das arithmetische Mittel aller ausgewerteten Intensitäten  $X$  des Gesamtsignals berechnet sich wie bekannt aus dem Integral der Verteilungsdichtefunktion  $p(x)$ :

$$\overline{X} = \int_{-\infty}^{\infty} x p(x) dx \quad \text{Gl. 1}$$

10

Eine Begrenzung der Integration bei einem bestimmten Wert  $x_G$  ermöglicht die Ermittlung des arithmetischen Mittelwertes über alle Werte  $X$ , die unter diesem Grenzwert liegen.

Dabei ist jedoch das Ergebnis mit der Häufigkeit  $P(x_G)$  zu wichten. Diese Häufigkeit entspricht dem Integral über  $p(x)$  bis zum Wert  $x_G$ .

15

$$\overline{X} = \int_{-\infty}^{x_G} x p(x) dx \Bigg/ \int_{-\infty}^{x_G} p(x) dx = \int_{-\infty}^{x_G} x p(x) dx \Bigg/ P(x_G) \quad \text{Gl. 2}$$

Der Intensitätsschwellwert  $x_G$  kann aus der Verteilungsfunktion  $P(x)$  abgeleitet werden. Im Beispiel nach Fig. 4 ist der Häufigkeitsschwellwert  $P(x_G)$  der Anteil von Sprachpausen im Gesamtsignal  $P_z = 0.58$ , dem der Intensitätsschwellwert  $x_G = 34.5 \text{ sone}$  zugeordnet ist. Das arithmetische Mittel aller Segmente mit einer Intensität, die geringer als  $x_G$  ist, berechnet sich nach Gl. 2, wobei  $x_G = 34.5 \text{ sone}$  gilt. Die Häufigkeit von 58% entspricht hier dem Wichtungswert  $P(x_G=34.5) = 0.58$ . Grafisch ist dieses Vorgehen in Fig. 5 dargestellt.

25 Wird nun wieder davon ausgegangen, dass die Intensitäten von Segmenten in Sprachpausen, die Intensitäten von Sprachsegmenten nicht übersteigen oder das Hintergrundgeräusch nur schwache zeitliche Fluktuationen aufweist, kann der berechnete arithmetische Mittelwert als Mittelwert der Intensität in Sprachpausen betrachtet werden.

### Vereinfachtes Verfahren zur Bestimmung des arithmetischen Mittelwertes

Ein vereinfachtes Verfahren zur Bestimmung des Mittelwertes über alle  $X$  geht von der Annahme aus, dass die relative Häufigkeitsverteilung der Intensitätswerte der

- 5 Signalsegmente im Bereich  $P(x) = 0$  bis zum Häufigkeitsschwellwert von Sprachpausen  $P_z$  durch eine gewichteten Normalverteilung  $G(x, \mu, \sigma^2)$  angenähert werden kann. Der Wert für die Verteilungsfunktion  $G(x, \mu, \sigma^2)$  für  $x \rightarrow \infty$  ist 1. Wie bekannt, entspricht der Wert  $x$ , bei dem gilt  $G(x, \mu, \sigma^2) = 0.5$ , dem arithmetischen Mittel über alle Einzelwerte  $X$ .
- 10 Gelingt eine Näherung der relativen Häufigkeitsverteilung  $P(x)$  im Bereich von  $P(x) = 0$  bis  $P_z$  mit einer gewichteten Normalverteilung  $\kappa P_z G(x, \mu, \sigma^2)$ , dann entspricht der arithmetische Mittelwert über  $X$  für die gewichtete Normalverteilung dem Wert  $x$  für den gilt:  $G(x, \mu, \sigma^2) = 0.5 \kappa P_z$ . Durch die Annahme, dass  $\kappa P_z G(x, \mu, \sigma^2)$  die Verteilung  $P(x)$  im Bereich von  $P(x) = 0$  bis  $P_z$  gut annähert und  $\kappa \geq 1$  ist, entspricht der gesuchte
- 15 arithmetische Mittelwert dem Wert  $x_\alpha$ , für den gilt  $P(x_\alpha) = 0.5 \kappa P_z$ .

Für den hier betrachteten Anwendungsfall von Sprache mit additivem Hintergrundgeräusch zeigen Werte für  $\kappa = 1 \dots 1.3$  gute Approximationsergebnisse. In Fig. 6 ist ein Beispiel für die Annäherung durch gewichtete Normalverteilungen gezeigt. Dabei wurde ein Wert

- 20  $\kappa = 1.1$  gewählt. Das Diagramm zeigt Sprache als Hintergrundgeräusch und hat einen Sprachpausenanteil von 58 %. Die starke zeitliche Fluktuation des Sprachhintergrundes lässt sich deutlich als flachere Steigung im Bereich  $N = 0 \dots 40$  sone erkennen. Der arithmetische Mittelwert, der aus der Normalverteilungsfunktion mit  $P(x_\alpha) = 0.5 \kappa P_z = 0.32$  abgeleitet wird, beträgt 20 sone.

25

Der Vorteil dieses vereinfachten Verfahrens ist die geringere Rechenintensität, da auf die Berechnung der Verteilungsdichte und deren Integration verzichtet werden kann. Es ist ebenfalls nicht notwendig, die Normalverteilungsfunktion  $\kappa P_z G(x, \mu, \sigma^2)$  exakt zu bestimmen, es genügt bereits die Festlegung von  $\kappa$ . Da  $P_z$  bekannt ist, wird der Mittelwert über alle  $X < x_\alpha$  als Wert  $x_\alpha$  bestimmt, bei dem gilt  $P(x_\alpha) = 0.5 \kappa P_z$ . Der arithmetische

30

Mittelwert über alle  $X$  bis  $x_c$  entspricht somit dem Intensitätswert, der einem Häufigkeitswert von  $0.5 * \kappa * \text{Anteil der Sprachpausen}$  am Gesamtsignal entspricht, d.h. der Intensität, die von einem Anteil von Segmenten von  $0.5 * \kappa * \text{Anteil der Sprachpausen}$  nicht überschritten wird.

5

#### ***Bestimmung weiterer statistischer Kennwerte***

Auch andere statistische Intensitätskennwerte können mit diesem Verfahren ermittelt werden. In Fig. 7 ist am Beispiel aus Fig. 4 demonstriert, wie aus der Funktion der Intensitätswert ermittelt werden kann, der von nur 20% der Sprachpausensegmente überschritten wird ( 20%-Perzentil-Lautheit).

10

Im angeführten Beispiel wird der Intensitätswert gesucht, der von 80% der Segmente in Sprachpausen unterschritten wird, d.h. gesucht wird der Abszissenwert, der für den Ordinatenwert  $P = 0.58 * 0.8 = 0.46$  gilt. Der Wert ist aufgrund des im Beispiel gewählten wenig schwankenden Störgeräusches nur wenig geringer als der Maximalwert.

15

#### ***Ausführungsbeispiel für die Bestimmung des arithmetischen Mittelwertes aus der Verteilungsdichtefunktion***

Das hier vorgestellte Ausführungsbeispiel des Verfahrens zur Intensitätsbestimmung von Hintergrundgeräuschen ermittelt den arithmetischen Mittelwert aller Lautheiten der Segmente, die unter einer bestimmten Häufigkeitsschwelle liegen. Diese Häufigkeitsschwelle entspricht dem Anteil an Sprachpausen im Signal und der errechnete arithmetische Mittelwert wird als mittlere Lautheit in Sprachpausen betrachtet. Dazu wird in diesem Ausführungsbeispiel die Verteilungsdichtefunktion benutzt.

20

Vorbedingung ist, dass beide Signale, d.h. das ungestörte Quellsprachsignal und das gestörte zu bewertende Signal, vollständig aufgezeichnet vorliegen.

25

Zunächst wird mittels einer geeigneten Schwelle anhand des Quellsprachsignals der Anteil an Sprachpausen  $P_z$  in diesem Signal bestimmt.

30

Der zweite Schritt ist die Berechnung der gewünschten Intensitätswerte für aufeinanderfolgende kurze Signalsegmente des zu bewertenden Sprachsignals. In diesem Ausführungsbeispiel werden die Lautheiten nach ISO532 in aufeinander folgenden Signalabschnitten von 16ms Länge berechnet. Die Verteilungsfunktion wird durch eine

- 5 Reihe von Einzelwerten (diskrete relative Häufigkeitsverteilung) angenähert. Diese Einzelwerte werden durch aufeinander folgende Indexe  $m$  bezeichnet. Die Reihe von Einzelwerten ist bei einem Maximalwert  $M$  begrenzt (z.B.:  $P_0 \dots P_M$ ). In der Auswertung wird jeder Einzelwert  $P_m$  – dessen Index die ermittelte Intensität  $X$  des ausgewerteten Signalsegments übersteigt – um den Zähler 1 erhöht. Nach Auswertung des gesamten
- 10 Signals werden alle Einzelwerte durch die Anzahl aller ausgewerteten Signal-semente dividiert. Jeder Einzelwert  $P_m$  enthält dann die relative Häufigkeit der Signal-semente, die eine Lautheit kleiner als der Wert des Indexes aufweisen.

Anhand des vorher ermittelten Anteils an Sprachpausen  $P_z$ , wird derjenige Häufigkeitswert

- 15  $P_s$  ermittelt, welcher die geringste absolute Differenz zu  $P_z$  besitzt. Der Index  $S$  dieses Einzelwertes  $P_s$  gibt die entsprechende Lautheit an, d.h. der Lautheit, die von einem Anteil  $P_s$  aller Segmente nicht überschritten wird. Zur Bestimmung des arithmetischen Mittels der Lautheiten aller Segmente, deren Lautheiten sich unter der vorgegebenen Häufigkeits-schwelle  $P_s$  befinden, ist als nächstes die Umwandlung der diskreten Häufigkeitsverteilung
- 20  $P_0 \dots P_M$  in eine diskrete Häufigkeitsdichte (Streifenhäufigkeit)  $p_0 \dots p_{M-1}$ , vorzunehmen. Dazu werden die Differenzen zweier aufeinanderfolgender Einzelwerte gebildet und als Wertefolge  $p_0 \dots p_{M-1}$ , abgelegt:

$$p_m = p_{m+1} - p_m$$

für alle  $m = 0 \dots M-1$

Gl. 3

25

Der Wert  $p_m$  enthält dann die relative Häufigkeit der Segmente, deren Lautheit sich zwischen  $m$  und  $m+1$  befindet. Der gesuchte arithmetische Mittelwert entspricht der gewichteten Summe über die Streifenhäufigkeit  $P_m$  bis  $m = S$ , d.h. der Lautheit, die von einem Anteil  $P_s$  aller Segmente nicht überschritten wird:

30

$$\tilde{N}_{av} = \sum_{m=0}^s \left( m + \frac{1}{2} \right) p_m \Bigg/ \sum_{m=0}^s p_m = \sum_{m=0}^s \left( m + \frac{1}{2} \right) p_m \Bigg/ P_s \quad \text{Gl. 4}$$

Der Korrekturwert  $\frac{1}{2}$  entspricht dem halben Abstand zweier aufeinander folgender Indexe.  
 Der Wert  $p_m$  enthält die relative Häufigkeit von Segmenten, deren Lautheiten sich zwischen  
 5  $m$  und  $m+1$  befinden. Der Erwartungswert aller hier erfassten Lautheiten ist, bei  
 angenommener Gleichverteilung der Lautheiten von  $m \dots m+1$ , daher  $m+0.5$ .

Das Verfahren liefert wie im Anwendungsfall beschrieben, eine diskrete Häufigkeitsverteilung mit einer Auflösung  $1 \text{ sone}$ , da der Index  $m$  ganzzahlig ist und die Lautheitswerte  
 10 direkt den entsprechenden Indexen zugeordnet werden. Um gegebenenfalls andere höhere oder verringerte Auflösungen zu erzielen ist der Lautheitswert vor Berechnung der relativen Häufigkeitsverteilung mit entsprechenden Faktoren zu multiplizieren.

Zur Demonstration der Messsicherheit des vorgestellten Verfahrens sind in Tabelle 1  
 15 Messwerte für verschiedene Signale und Hintergrundgeräusche aufgeführt. Es wurde Sprachsignale von 32 s Länge und verschiedenem Anteil an Sprachpausen (35%, 58% und 91%) jeweils mit verschiedenen Geräuschen gemischt. Als Geräusche wurde zunächst weißes Rauschen mit verschiedenen Sprach-Geräusch-Abständen benutzt. Des weiteren wurde auch kontinuierlich gesprochene Sprache sowie zwei Geräusche aus realen  
 20 akustischen Umgebungen (Straße und Büro) eingesetzt.

Vor Berechnung der Häufigkeitsverteilung wird eine Multiplikation aller Lautheitswerte mit dem Faktor 2 durchgeführt, um die Auflösung der Darstellung bei Benutzung ganzzahliger Indexe zu erhöhen. Dies entspricht dann einer Lautheitsstufung bei ganzzahligen Indexen  
 25 von  $0.5 \text{ sone}$ . Mit einer Begrenzung der Häufigkeitsverteilungsfunktion bei  $P_{2m}$ , können so Lautheiten von  $0 \dots 100 \text{ sone}$  in Schritten von  $0.5 \text{ sone}$  abgebildet werden. Es ist aber zu beachten, dass dieser Faktor als Divisor zur Korrektur auf alle Ergebnisse angewendet wird muss. Im hier gewählten Ausführungsbeispiel bedeutet dies, dass der errechnete arithmetische Mittelwert durch 2 zu teilen ist.

Erläuterungen zu Tabelle 1: Der Sprach-Geräusch-Abstand dient lediglich zur Information; Grundlage bildet der Abstand des mittleren Effektivpegels bei Sprachaktivität zum mittleren Effektivpegel des Hintergrundgeräusches. Der mittlere Lautheitswert (Zielwert) wurde in einer Referenzmessung bestimmt, bei der die Sprachpausen manuell markiert und in Segmenten zu 16 ms ausgewertet wurden. Die berechneten Standardabweichungen beziehen sich auf die derart gemessenen Referenz-Lautheiten und geben Information über die Stärke der auftretenden Fluktuationen. Die Messwerte in Spalte 5 wurden mit dem in diesem Ausführungsbeispiel beschriebenen Verfahren ermittelt.

Geräusch	SNR	mittlere Lautheit (sone) Zielwert	Standardabweichung der Segment-Lautheiten	mittlere Lautheit (sone) gemessen mit beschriebenen Verfahren	Abweichung (Messfehler) abs. / rel.
<b>Pausenanteil des Sprachsignals 91%</b>					
weißes Rauschen	6 dB	41.4	1.55	42.0	0.6 / 1.4%
weißes Rauschen	10 dB	32.3	1.22	32.6	0.3 / 0.9%
weißes Rauschen	16 dB	22.2	0.87	22.3	0.1 / 0.4%
Sprache	6 dB	21.3	11.7	20.6	-0.7 / -3.3%
Sprache	10 dB	16.5	9.16	16.2	-0.3 / -1.8%
Sprache	16 dB	11.2	6.21	11.3	0.1 / 0.9%
Straßengeräusch	10 dB	26.0	3.22	26.2	0.2 / 0.8%
Bürogeräusch	10 dB	26.3	2.78	26.6	0.3 / 1.1%
<b>Pausenanteil des Sprachsignals 58%</b>					
weißes Rauschen	6 dB	41.3	1.55	44.8	3.5 / 8.5%
weißes Rauschen	10 dB	32.3	1.22	34.2	1.9 / 6.0%
weißes Rauschen	16 dB	22.1	0.87	22.6	0.5 / 2.2%
Sprache	6 dB	20.7	11.7	19.0	-1.7 / -8.2%
Sprache	10 dB	16.0	9.16	15.4	-0.6 / -3.8%
Sprache	16 dB	10.7	6.21	10.8	0.1 / 0.9%
Straßengeräusch	10 dB	26.1	3.22	27.0	0.9 / 3.4%
Bürogeräusch	10 dB	26.3	2.78	27.3	1.0 / 3.8%
<b>Pausenanteil des Sprachsignals 35%</b>					
weißes Rauschen	6 dB	41.3	1.55	46.1	4.8 / 11.6%
weißes Rauschen	10 dB	32.3	1.22	35.6	3.3 / 10.2%
weißes Rauschen	16 dB	22.1	0.87	23.3	1.2 / 5.4%

Sprache	6 dB	<b>20.0</b>	<i>11.22</i>	<b>17.6</b>	-2.4 / -12%
Sprache	10 dB	<b>15.6</b>	<i>8.7</i>	<b>15.0</b>	-0.6 / -3.8%
Sprache	16 dB	<b>10.9</b>	<i>5.93</i>	<b>11.8</b>	0.9 / 8.3%
Straßengeräusch	10 dB	<b>26.1</b>	<i>3.22</i>	<b>27.3</b>	1.2 / 4.6%
Bürogeräusch	10 dB	<b>26.3</b>	<i>2.78</i>	<b>27.9</b>	1.6 / 6.1%

Tabelle I

Zunächst ist festzustellen, dass die Messsicherheit mit zunehmenden Pausenanteil im zu bewertenden Signal zunimmt. Eine Zunahme der Messsicherheit ist ebenfalls bei sinkender 5 Geräuschintensität sowie geringerer zeitlicher Fluktuation des Hintergrundgeräusches festzustellen. Ausgehend von einem typischen Anteil an Sprachpausen in einer Telefonkommunikation von  $P_z > 50\%$  sind die mit dem vorgestellten Verfahren erreichten Messwerte selbst bei stärkeren Fluktuationen im Hintergrundgeräusch (z.B. Sprache) zufriedenstellend.

10

#### Ausführungsbeispiel für die Bestimmung des arithmetischen Mittelwertes mit vereinfachtem Verfahren

Dieses spezielle Ausführungsbeispiel zeigt eine Anwendung des beschriebenen vereinfachten Verfahren zur Bestimmung des arithmetischen Mittels unter Nutzung einer 15 gewichteten Normalverteilung.

Das vereinfachte Verfahren verzichtet auf die Berechnung der Streifenhäufigkeit und leitet einen Schätzwert für das arithmetische Mittel der Lautheiten aller Segmente, deren Lautheiten sich unter der vorgegebenen Häufigkeitsschwelle  $P_z$  befinden, direkt aus der 20 relativen Häufigkeitsverteilung  $P_m$  ab. Wie beschrieben muss lediglich der Wert  $\kappa$  für die Schätzung festgelegt werden.

In diesem Ausführungsbeispiel wird mit  $\kappa = 1.1$  definiert. Der Schätzwert entspricht dann dem Lautheitswert, der von einem Anteil von  $0.5 * 1.1 * P_z$  aller ausgewerteten Segmente 25 nicht überschritten wird. Im Ausführungsbeispiel entspricht dieser Schätzwert des arithmetischen Mittels der Lautheiten, dem Index  $m$  des Häufigkeitswertes, welcher die

geringste absolute Differenz zu  $0.55 P_z$  besitzt. In Tabelle 2 sind die Messwerte aufgeführt, die mit diesem vereinfachten Verfahren gewonnen worden. Auch hier wurden zur Erhöhung der Auflösung auf  $0.5 \text{ sone}$  alle Lautheitswerte vor Berechnung der Häufigkeitsverteilung mit dem Faktor 2 multipliziert und die Ergebnisse entsprechend korrigiert.

5

Geräusch	SNR	mittlere Lautheit (sone) Zielwert	Standardabweichung der Segment-Lautheiten	mittlere Lautheit (sone) gemessen mit vereinfachtem Verfahren	Abweichung (Messfehler) abs. / rel.
<b>Pausenanteil des Sprachsignals 91%</b>					
weißes Rauschen	6 dB	<b>41.4</b>	<i>1.55</i>	<b>41.5</b>	0.1 / 0.2%
weißes Rauschen	10 dB	<b>32.3</b>	<i>1.22</i>	<b>32.5</b>	0.2 / 0.6%
weißes Rauschen	16 dB	<b>22.2</b>	<i>0.87</i>	<b>22.5</b>	0.3 / 1.3%
Sprache	6 dB	<b>21.3</b>	<i>11.7</i>	<b>20.5</b>	-0.8 / -3.8%
Sprache	10 dB	<b>16.5</b>	<i>9.16</i>	<b>16.5</b>	0.0 / 0.0%
Sprache	16 dB	<b>11.2</b>	<i>6.21</i>	<b>11.0</b>	-0.2 / 1.8%
Straßengeräusch	10 dB	<b>26.0</b>	<i>3.22</i>	<b>26.0</b>	0.0 / 0.0%
Bürogeräusch	10 dB	<b>26.3</b>	<i>2.78</i>	<b>26.5</b>	0.2 / 0.6%
<b>Pausenanteil des Sprachsignals 58%</b>					
weißes Rauschen	6 dB	<b>41.3</b>	<i>1.55</i>	<b>41.50</b>	0.2 / 0.5%
weißes Rauschen	10 dB	<b>32.3</b>	<i>1.22</i>	<b>32.5</b>	0.2 / 0.6%
weißes Rauschen	16 dB	<b>22.1</b>	<i>0.87</i>	<b>22.5</b>	0.4 / 1.8%
Sprache	6 dB	<b>20.7</b>	<i>11.7</i>	<b>20.0</b>	-0.7 / -3.4%
Sprache	10 dB	<b>16.0</b>	<i>9.16</i>	<b>16.0</b>	0.0 / 0.0%
Sprache	16 dB	<b>10.7</b>	<i>6.21</i>	<b>11.0</b>	0.3 / 2.8%
Straßengeräusch	10 dB	<b>26.1</b>	<i>3.22</i>	<b>26.0</b>	-0.1 / -0.4%
Bürogeräusch	10 dB	<b>26.3</b>	<i>2.78</i>	<b>26.5</b>	0.2 / 0.8%
<b>Pausenanteil des Sprachsignals 35%</b>					
weißes Rauschen	6 dB	<b>41.3</b>	<i>1.55</i>	<b>41.0</b>	-0.3 / 0.7%
weißes Rauschen	10 dB	<b>32.3</b>	<i>1.22</i>	<b>32.5</b>	0.2 / 0.6%
weißes Rauschen	16 dB	<b>22.1</b>	<i>0.87</i>	<b>22.5</b>	0.4 / 1.8%
Sprache	6 dB	<b>20.0</b>	<i>11.22</i>	<b>19.0</b>	-1.0 / -5%
Sprache	10 dB	<b>15.6</b>	<i>8.7</i>	<b>15.5</b>	-0.1 / -0.6%
Sprache	16 dB	<b>10.9</b>	<i>5.93</i>	<b>11.5</b>	0.6 / 5.5%

Straßengeräusch	10 dB	<b>26.1</b>	<b>3.22</b>	<b>25.5</b>	-0.6 / -1.4%
Bürogeräusch	10 dB	<b>26.3</b>	<b>2.78</b>	<b>26.5</b>	0.2 / 0.8%

Tabelle 2

Das vereinfachte Verfahren spart nicht nur Rechenzeit sondern liefert in den ausgewerteten Beispielen Messwerte mit einer deutlich höheren Genauigkeit im Vergleich zu den Werten aus Tabelle 1. Da als Schätzwert direkt der Index  $m$  benutzt wird, ist die Genauigkeit der 5 Schätzung auf die Auflösung der relativen diskreten Häufigkeitsverteilung (hier: *0.5 sone*) begrenzt.

Mit dem beschriebenen vereinfachten Messverfahren werden auch bei Geräuschen mit stärkerer Fluktuation gute Messwerte erzielt. Bei den gewählten Sprach-Geräusch- 10 Abständen von *6dB* kann auch nicht mehr davon ausgegangen werden, dass alle Lautheiten in Sprachpausen eine geringere Lautheit als Sprachsegmente aufweisen. Trotzdem sind die Messwerte kaum verfälscht wurden. Das beschriebene vereinfachte Verfahren eignet sich zudem auch für Signale mit geringerem Pausenanteil.

15 **Ausführungsbeispiel für die Bestimmung von Perzentil-Lautheiten aus der relativen Häufigkeitsverteilung**

Die Perzentil-Lautheit aller Segmente, die unter einer bestimmten Häufigkeitsschwelle  $P_z$  liegen, kann durch Multiplikation dieser relativen Häufigkeit  $P_z$  mit einem Wert  $I$  - 20 *Perzentilwert* erfolgen (z.B. 10%-Perzentil-Lautheit:  $P_{z,10\%} = 0.9 * P_z$ ). Der ganzzahlige Index  $m$  des Häufigkeitswertes  $P_m$ , welcher die geringste absolute Differenz zu  $P_{z,10\%}$  besitzt, liefert den gesuchten Perzentil-Lautheitswert.

In Tabelle 3 sind für die bereits in den Tabellen 1 und 2 aufgeführten Beispiele die 10%-Perzentil-Lautheiten aufgeführt und werden mit einem manuell bestimmten Referenzwert 25 verglichen.

Geräusch	SNR	10%-Perzentil-Lautheit (sone) Zielwert	Standard-abweichung der Segment-Lautheiten	10%-Perzentil-Lautheit (sone) gemessen über Häufigkeitsverteilung	Abweichung (Messfehler) abs. / rel.
<b>Pausenanteil des Sprachsignals 91%</b>					
weißes Rauschen	6 dB	<b>42.5</b>	<i>1.55</i>	<b>43.0</b>	0.5 / 1.2%
weißes Rauschen	10 dB	<b>33.0</b>	<i>1.22</i>	<b>34.0</b>	1.0 / 3.0%
weißes Rauschen	16 dB	<b>22.5</b>	<i>0.87</i>	<b>23.5</b>	1.0 / 4.4%
Sprache	6 dB	<b>37.0</b>	<i>11.7</i>	<b>34.5</b>	-2.5 / -6.8%
Sprache	10 dB	<b>28.5</b>	<i>9.16</i>	<b>27.5</b>	-1.0 / -3.5%
Sprache	16 dB	<b>19.0</b>	<i>6.21</i>	<b>19.5</b>	0.5 / 2.6%
Straßengeräusch	10 dB	<b>29.5</b>	<i>3.22</i>	<b>30.0</b>	0.5 / 1.7%
Bürogeräusch	10 dB	<b>29.0</b>	<i>2.78</i>	<b>29.5</b>	0.5 / 1.7%
<b>Pausenanteil des Sprachsignals 58%</b>					
weißes Rauschen	6 dB	<b>42.5</b>	<i>1.55</i>	<b>42.5</b>	0.0 / 0.0%
weißes Rauschen	10 dB	<b>33.0</b>	<i>1.22</i>	<b>33.5</b>	0.5 / 1.5%
weißes Rauschen	16 dB	<b>22.5</b>	<i>0.87</i>	<b>23.0</b>	0.5 / 2.2%
Sprache	6 dB	<b>36.0</b>	<i>11.7</i>	<b>29.0</b>	-7.0 / -19%
Sprache	10 dB	<b>28.5</b>	<i>9.16</i>	<b>24.5</b>	-4.0 / -14%
Sprache	16 dB	<b>19.0</b>	<i>6.21</i>	<b>18.0</b>	-1.0 / -5.3%
Straßengeräusch	10 dB	<b>30.0</b>	<i>3.22</i>	<b>29.0</b>	-1.0 / -3.3%
Bürogeräusch	10 dB	<b>29.0</b>	<i>2.78</i>	<b>28.5</b>	-0.5 / -1.6%
<b>Pausenanteil des Sprachsignals 35%</b>					
weißes Rauschen	6 dB	<b>42.5</b>	<i>1.55</i>	<b>42.5</b>	0.0 / 0.0%
weißes Rauschen	10 dB	<b>33.0</b>	<i>1.22</i>	<b>33.5</b>	0.5 / 1.5%
weißes Rauschen	16 dB	<b>22.5</b>	<i>0.87</i>	<b>23.5</b>	1.0 / 2.2%
Sprache	6 dB	<b>35.5</b>	<i>11.22</i>	<b>24.0</b>	-11.5 / -33%
Sprache	10 dB	<b>27.5</b>	<i>8.7</i>	<b>21.0</b>	-6.5 / -24%
Sprache	16 dB	<b>19.0</b>	<i>5.93</i>	<b>17.5</b>	-1.5 / -7.9%
Straßengeräusch	10 dB	<b>29.5</b>	<i>3.22</i>	<b>28.0</b>	-1.5 / -4.8%
Bürogeräusch	10 dB	<b>29.0</b>	<i>2.78</i>	<b>28.5</b>	-0.5 / -1.6%

Tabelle 3

Die Messwerte zeigen eine gute Abschätzung der Perzentil-Lautheit für Hintergrundgeräusche mit schwacher Fluktuation, für Sprache werden – vor allem bei geringem Pausenanteil – nur unzureichende Genauigkeiten erzielt. Lediglich bei höheren Sprach-Geräusch-Abständen sind die Ergebnisse brauchbar bis gut.

**Patentansprüche (4)**

1. Verfahren zur Bestimmung von Intensitätskennwerten von Hintergrundgeräuschen in Sprachpausen von Sprachsignalen, von denen das ungestörte Quellsprachsignal und das gestörte Sprachsignal aufgezeichnet zur Verfügung stehen und aus dem ungestörten Quellsprachsignal der Anteil Sprachpausen im Gesamtsignal nach bekannten Methoden ermittelt wird und das gestörte Sprachsignal in kurze aufeinanderfolgende Signalelemente unterteilt wird und für jedes Signalelement ein Intensitätswert bestimmt wird, **d a d u r c h g e k e n n z e i c h n e t**,  
dass aus den Intensitätswerten der einzelnen Signalelemente des gestörten Sprachsignals die kumulative relative Häufigkeitsverteilung (1) gebildet wird,  
dass der ermittelte Anteil an Sprachpausen im Quellsprachsignal als Häufigkeitsschwelle definiert wird und die Häufigkeitsschwelle auf das gestörte Sprachsignal angewendet wird,  
dass aus der Häufigkeitsverteilung der Intensitätswerte der Signalsegmente der definierten Häufigkeitsschwelle (2) entsprechende Intensitätsschwellwert (3) ermittelt wird,  
dass alle Signalsegmente mit einem geringeren Intensitätswert als dem des Intensitätsschwellwertes zu den Sprachpausen gehörend bewertet werden,  
dass die Verteilungsfunktion für die Intensitätswerte der Signalsegmente in dem Bereich unterhalb des Intensitätsschwellwertes die Häufigkeitsverteilung für die Intensitätswerte in den Sprachpausen (4) darstellt, und dass dieser Bereich der Verteilungsfunktion für die Ermittlung von Intensitätskennwerten von den Hintergrundgeräuschen in den Sprachpausen genutzt werden kann.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass als Intensitätskennwert von den Hintergrundgeräuschen in den Sprachpausen der arithmetische Mittelwert der Intensitätswerte der Signalelemente in den Sprachpausen bestimmt wird, und dass der arithmetische Mittelwert berechnet wird, indem aus der Häufigkeitsverteilung die Verteilungsdichte abgeleitet wird und durch eine nachfolgende Integration über die

Verteilungsdichte im Bereich unterhalb des Intensitätsschwellwertes der arithmetische Mittelwert der Intensitätswerte in den Sprachpausen ermittelt wird.

3. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass als Intensitätskennwert von den Hintergrundgeräuschen in den Sprachpausen der arithmetische Mittelwert der Intensitätswerte der Signalelemente in den Sprachpausen bestimmt wird, und dass der arithmetische Mittelwert aus der Häufigkeitsverteilung ermittelt wird, indem die Intensitätsverteilung im Bereich unterhalb des Intensitätsschwellwertes durch eine mit einem Faktor gewichtete Normalverteilung angenähert wird und für die Berechnung des arithmetischen Mittelwertes der Intensitätsschwellwert mit 0,5 und dem Wichtungsfaktor multipliziert wird.
4. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass als Intensitätskennwerte von den Hintergrundgeräuschen in den Sprachpausen Perzentil-Kennwerte bestimmt werden können, dass die Perzentil-Kennwerte aus der Häufigkeitsverteilung ermittelt werden können, indem der vorgegebene Perzentilwert von 100 Prozent subtrahiert wird, die Differenz mit dem Häufigkeitsschwellwert multipliziert wird und für den sich ergebenden Häufigkeitswert der diesem Wert entsprechende Intensitätswert als Perzentil-Kennwert aus der Verteilfunktion bestimmt wird.

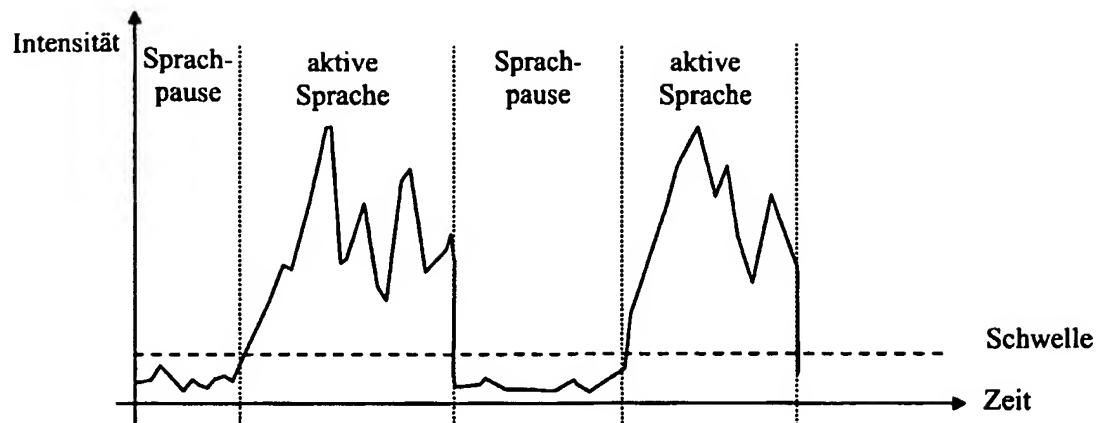


Fig. 1

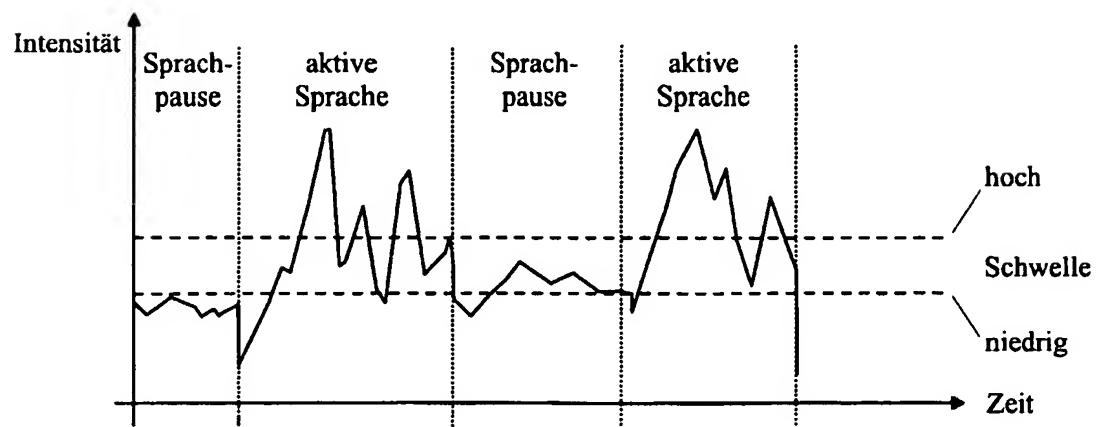
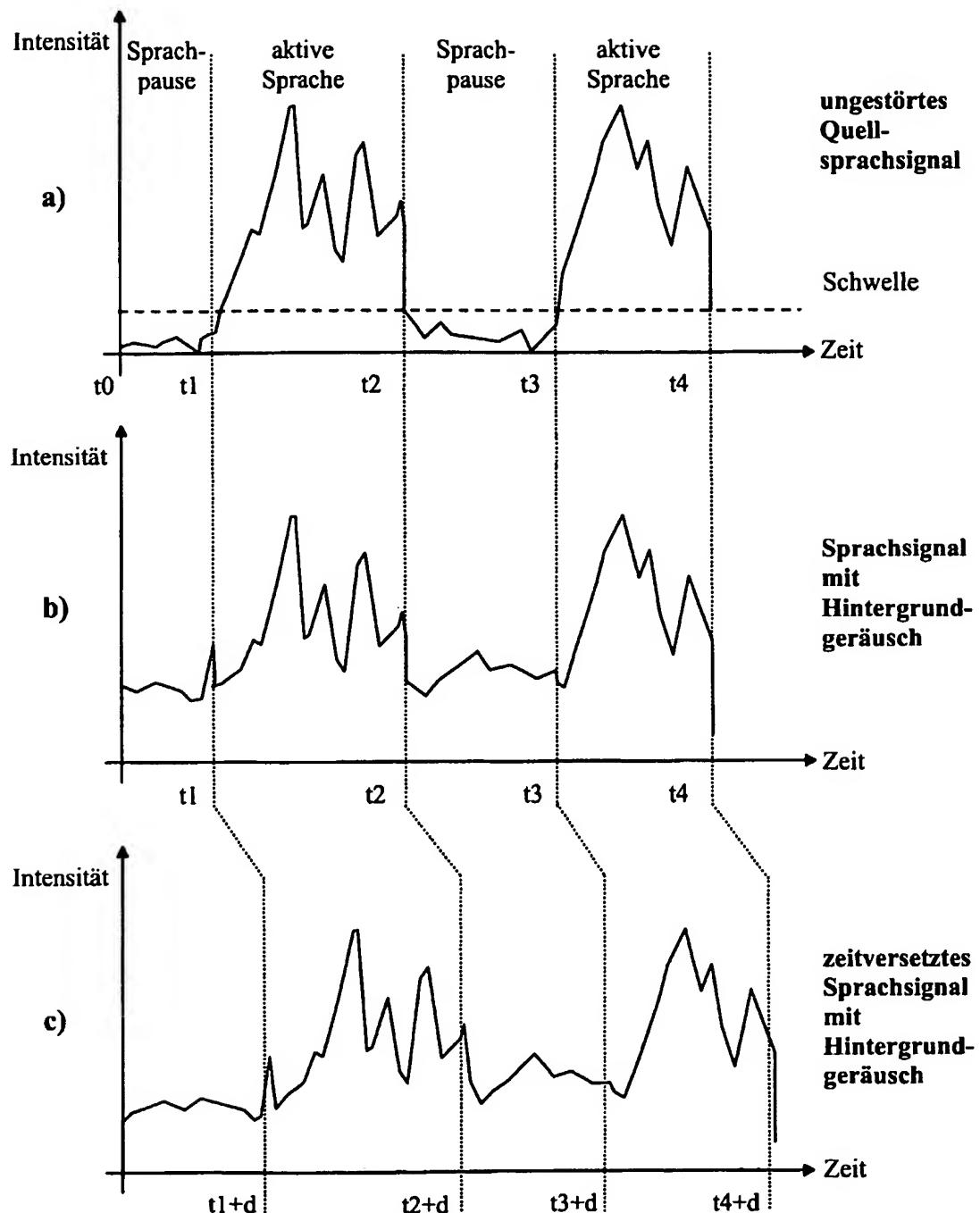


Fig. 2

**Fig. 3**

3/5

Relative Häufigkeitsverteilung der Kurzzeit-Intensitäten  
des gestörten Sprachsignals

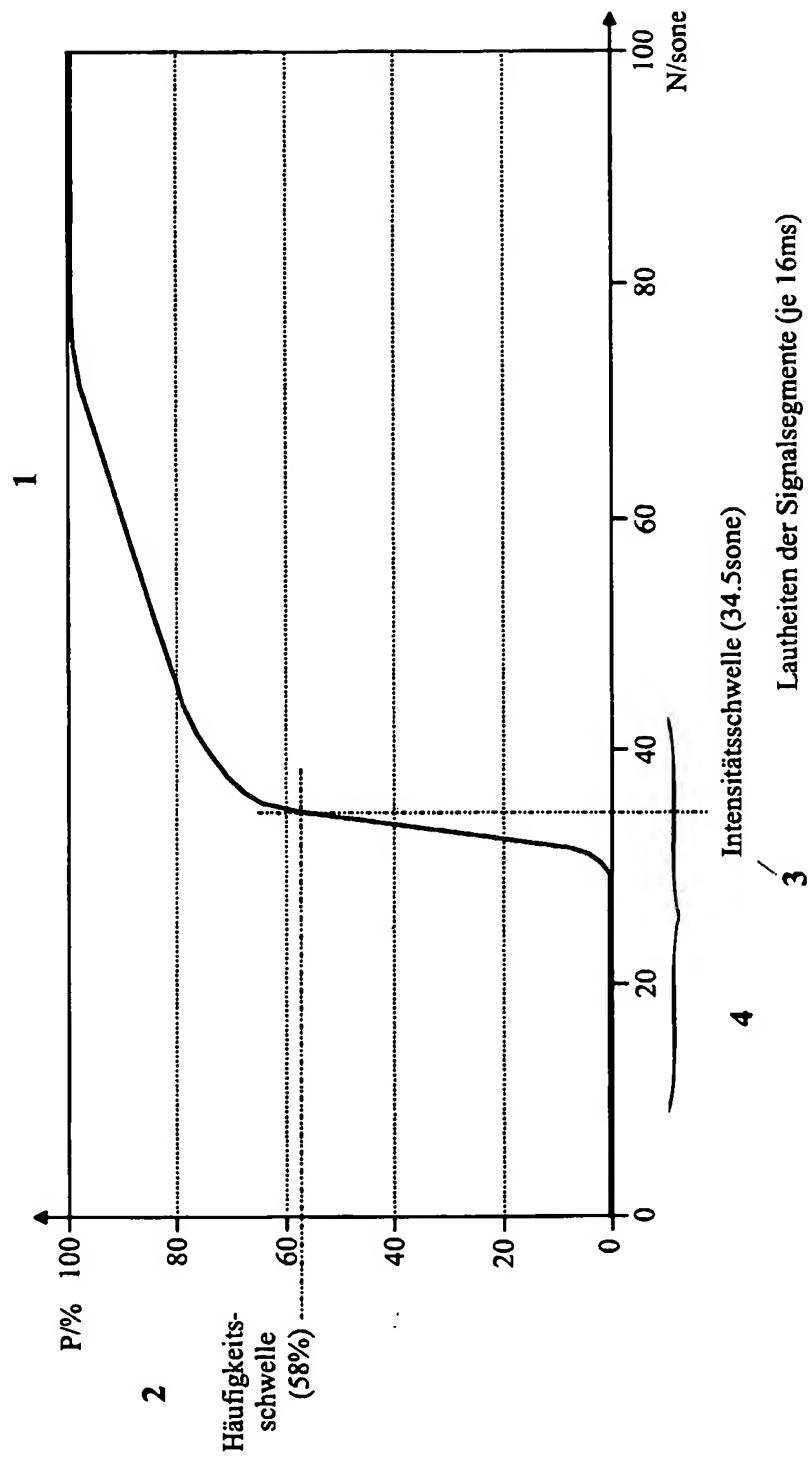
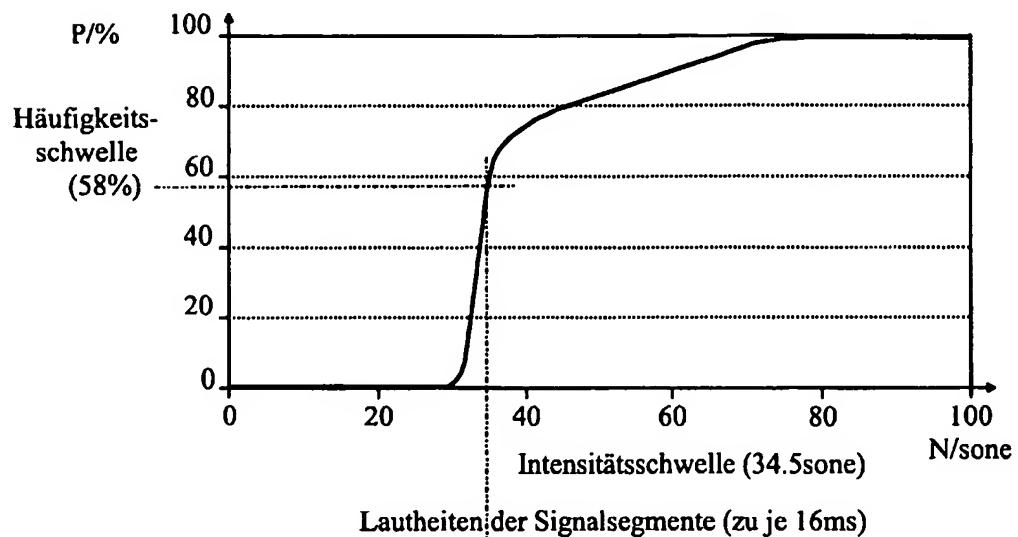
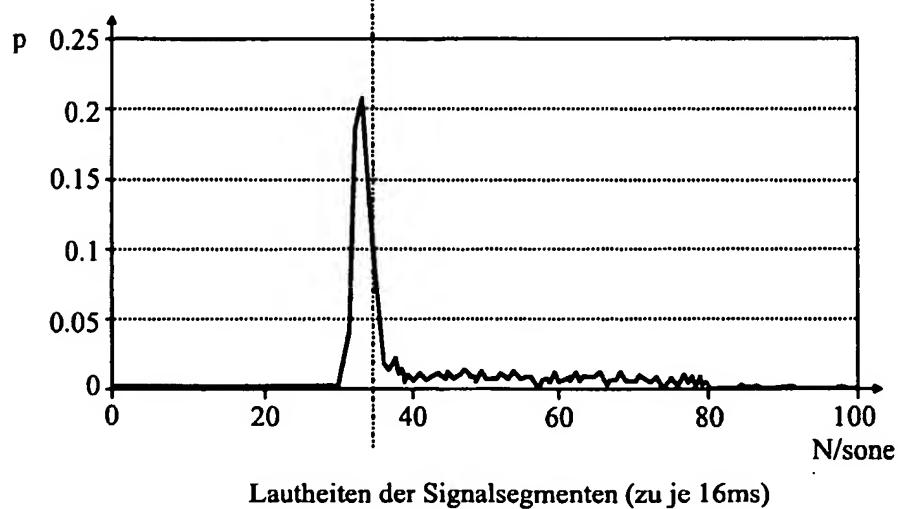


Fig. 4

### Häufigkeitsverteilung von Lautheiten in Signalsegmenten

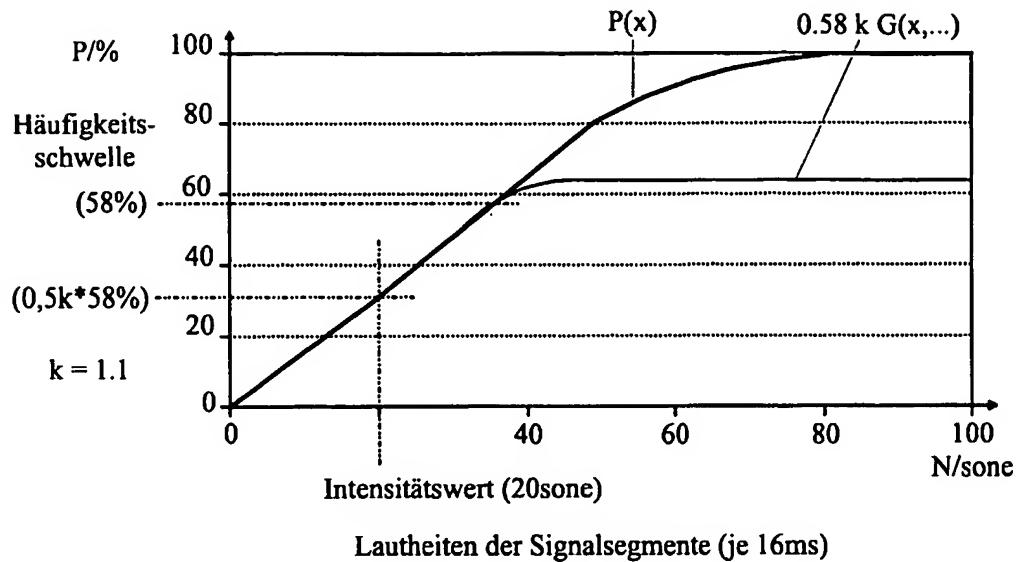
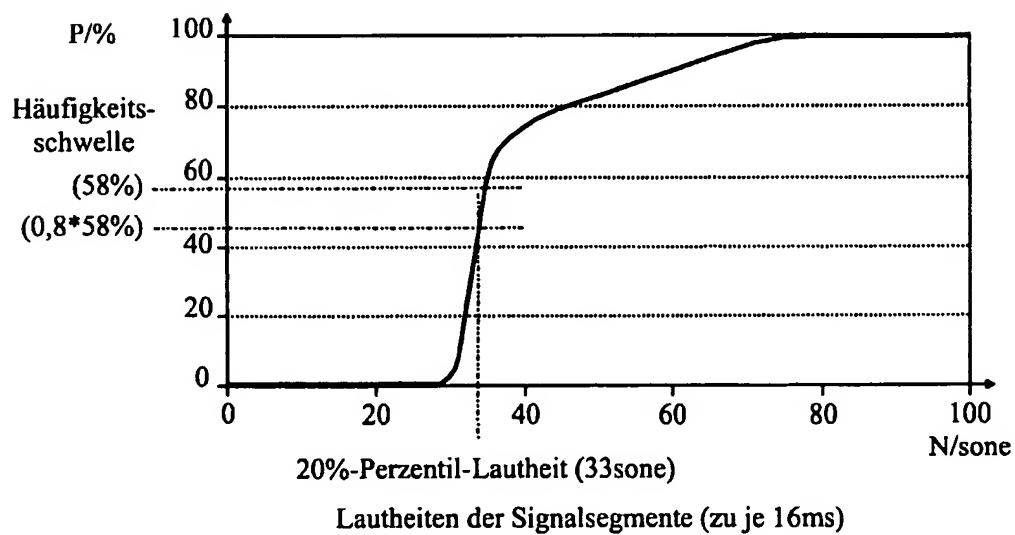


### Verteilungsdichte von Lautheiten in Signalsegmenten



**Fig.5**

5/5

**Fig. 6****Fig. 7**

## INTERNATIONAL SEARCH REPORT

tional Application No

PCT/DE 02/01200

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 7 G10L19/00 G10L11/02

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 G10L H04M

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

WPI Data, PAJ, EPO-Internal, INSPEC

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 6 044 342 A (AOKI MAKOTO ET AL) 28 March 2000 (2000-03-28) abstract column 3, line 37 - line 46 column 6, line 6 - line 32 column 6, line 37 - line 49 ---	1
A	US 5 598 466 A (GRAUMANN DAVID L) 28 January 1997 (1997-01-28) column 5, line 61 -column 6, line 61 ---	1
A	WO 00 52683 A (PANASONIC TECHNOLOGIES INC) 8 September 2000 (2000-09-08) page 7, line 19 - line 29 ---	1
A	US 4 811 404 A (VILMUR RICHARD J ET AL) 7 March 1989 (1989-03-07) column 3, line 4 - line 34 ---	1

 Further documents are listed in the continuation of box C. Patent family members are listed in annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the International filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the International filing date but later than the priority date claimed

- \*T\* later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*Z\* document member of the same patent family

Date of the actual completion of the international search

7 August 2002

Date of mailing of the international search report

14/08/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Ramos Sánchez, U

## INTERNATIONAL SEARCH REPORT

Information on patent family members

National Application No

PCT/DE 02/01200

Patent document cited in search report		Publication date		Patent family member(s)	Publication date
US 6044342	A	28-03-2000	JP JP	3255584 B2 10210075 A	12-02-2002 07-08-1998
US 5598466	A	28-01-1997	AU EP WO US US	6505196 A 0847645 A1 9708882 A1 5737407 A 6175634 B1	19-03-1997 17-06-1998 06-03-1997 07-04-1998 16-01-2001
WO 0052683	A	08-09-2000	US EP WO	6327564 B1 1163666 A1 0052683 A1	04-12-2001 19-12-2001 08-09-2000
US 4811404	A	07-03-1989	DE DE EP JP JP KR WO	3856280 D1 3856280 T2 0380563 A1 2995737 B2 3500347 T 9700789 B1 8903141 A1	21-01-1999 12-08-1999 08-08-1990 27-12-1999 24-01-1991 20-01-1997 06-04-1989

# INTERNATIONÄLER RECHERCHENBERICHT

tionales Aktenzeichen

PC1/DE 02/01200

**A. KLASIFIZIERUNG DES ANMELDUNGSGEGENSTANDES**  
IPK 7 G10L19/00 G10L11/02

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

**B. RECHERCHIERTE GEBIETE**

Recherchierte Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)  
IPK 7 G10L H04M

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der Internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

WPI Data, PAJ, EPO-Internal, INSPEC

**C. ALS WESENTLICH ANGESEHENE UNTERLAGEN**

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	US 6 044 342 A (AOKI MAKOTO ET AL) 28. März 2000 (2000-03-28) Zusammenfassung Spalte 3, Zeile 37 - Zeile 46 Spalte 6, Zeile 6 - Zeile 32 Spalte 6, Zeile 37 - Zeile 49 ---	1
A	US 5 598 466 A (GRAUMANN DAVID L) 28. Januar 1997 (1997-01-28) Spalte 5, Zeile 61 - Spalte 6, Zeile 61 ---	1
A	WO 00 52683 A (PANASONIC TECHNOLOGIES INC) 8. September 2000 (2000-09-08) Seite 7, Zeile 19 - Zeile 29 ---	1
A	US 4 811 404 A (VILMUR RICHARD J ET AL) 7. März 1989 (1989-03-07) Spalte 3, Zeile 4 - Zeile 34 ----	1

Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

Siehe Anhang Patentfamilie

- \* Besondere Kategorien von angegebenen Veröffentlichungen :
- \*'A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist
- \*'E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist
- \*'L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)
- \*'O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht
- \*'P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist
- \*'T" Später Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist
- \*'X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden
- \*'Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist
- \*'g" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

Absendedatum des internationalen Recherchenberichts

7. August 2002

14/08/2002

Name und Postanschrift der Internationalen Recherchenbehörde  
Europäisches Patentamt, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Bevollmächtigter Bediensteter

Ramos Sánchez, U

**INTERNATIONALER RECHERCHENBERICHT**

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

ionales Aktenzeichen

PCT/DE 02/01200

Im Recherchenbericht angeführtes Patentdokument		Datum der Veröffentlichung		Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
US 6044342	A	28-03-2000	JP JP	3255584 B2 10210075 A	12-02-2002 07-08-1998
US 5598466	A	28-01-1997	AU EP WO US US	6505196 A 0847645 A1 9708882 A1 5737407 A 6175634 B1	19-03-1997 17-06-1998 06-03-1997 07-04-1998 16-01-2001
WO 0052683	A	08-09-2000	US EP WO	6327564 B1 1163666 A1 0052683 A1	04-12-2001 19-12-2001 08-09-2000
US 4811404	A	07-03-1989	DE DE EP JP JP KR WO	3856280 D1 3856280 T2 0380563 A1 2995737 B2 3500347 T 9700789 B1 8903141 A1	21-01-1999 12-08-1999 08-08-1990 27-12-1999 24-01-1991 20-01-1997 06-04-1989

This Page is inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT OR DRAWING
- BLURED OR ILLEGIBLE TEXT OR DRAWING
- SKEWED/SLANTED IMAGES
- COLORED OR BLACK AND WHITE PHOTOGRAPHS
- GRAY SCALE DOCUMENTS
- LINES OR MARKS ON ORIGINAL DOCUMENT
- REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**  
**As rescanning documents *will not* correct images problems checked, please do not report the problems to the IFW Image Problem Mailbox**